



Development of dam inflow simulation technique coupled with rainfall simulation and rainfall-runoff model

Kim, Tae-Jeong^a · So, Byung-Jin^a · Ryou, Min-Suk^b · Kwon, Hyun-Han^{a*}

^aDepartment of Civil Engineering, Chonbuk National University, Jeonju, Korea

^bK-water, 200, Sintanjin-ro, Daedeok-gu, Daejeon, 306, Korea

Paper number: 16-010

Received: 29 January 2016; Revised: 16 February 2016 / 23 February 2016; Accepted: 23 February 2016

Abstract

Generally, a natural river discharge is highly regulated by the hydraulic structures, and the regulated flow is substantially different from natural inflow characteristics for the use of water resources planning. The natural inflow data are necessarily required for hydrologic analysis and water resources planning. This study aimed to develop an integrated model for more reliable simulation of daily dam inflow. First, a piecewise Kernel-Pareto distribution was used for rainfall simulation model, which can more effectively reproduce the low order moments (e.g. mean and median) as well as the extremes. Second, a Bayesian Markov Chain Monte Carlo scheme was applied for the SAC-SMA rainfall-runoff model that is able to quantitatively assess uncertainties associated with model parameters. It was confirmed that the proposed modeling scheme is capable of reproducing the underlying statistical properties of discharge, and can be further used to provide a set of plausible scenarios for water budget analysis in water resources planning.

Keywords: Discharge, Rainfall Simulation, Bayesian, Rainfall-Runoff Model

강우모의기법과 강우-유출 모형을 연계한 댐 유입량 자료 생성기법 개발

김태정^a · 소병진^a · 유민석^b · 권현한^{a*}

^a전북대학교 토목공학과, ^b한국수자원공사 댐·유역관리처

요 지

일반적으로 하천의 유량은 댐과 같은 수공구조물에 의해 조정된 유량으로 수자원계획을 위해서 필요한 자연유량과는 차이가 크다. 수자원계획을 수립함에 있어 자연 유입량 정보는 댐 운영과 수문분석을 위한 필수적인 정보이다. 본 연구에서는 댐 유역 일유입량 모의기법을 위한 통합 모형을 개발하였다. 첫째, 장기 강우-유출 모형의 입력강우자료로 사용하기 위하여 평균 및 중앙값과 같은 통계적 모멘트를 효과적으로 재현하고 극치 강우량 재현에 유리한 불연속 Kernel-Pareto 확률분포 기반의 강우모의기법을 통하여 강우모의를 수행하였다. 둘째, SAC-SMA 장기 강우-유출 모형의 매개변수를 Bayesian MCMC 기법을 통하여 최적화하여 산정된 매개변수의 사후분포를 활용하여 댐 유입량 시나리오 도출하였다. 댐 유역을 대상으로 개발된 모형을 평가한 결과 자연유량과 통계적으로 유사한 특성을 가지는 시나리오를 생성할 수 있었으며, 물수지 분석 등과 같은 수자원계획을 위한 시나리오로 활용이 가능할 것으로 판단된다.

핵심용어: 유량, 강우모의, Bayesian, 강우-유출 모형

1. 서 론

최근 한반도는 극심한 가뭄으로 인하여 수자원관리 측면

에서 많은 어려움이 발생되고 있으며 전 세계적으로 대형 가뭄이 계속 발생하고 있다(Fallah and Cubasch, 2015). 태국은 2014년 발생한 극심한 가뭄으로 인한 쌀 생산 감소로 국가경제에 큰 타격을 입을 것으로 예상되고 있다. 미국 캘리포니아 주는 최근 심각한 가뭄으로 인하여 주민들에게 ‘강제 절수령’을 선포하여 주민들에게 많은 불편함을 유발하여 수자원관리

*Corresponding Author. Tel: +82-63-270-2426
E-mail: hkwon@jbnu.ac.kr (H.-H. Kwon)

의 심각성이 부각되고 있다(Wang et al., 2014; Diffenbaugh et al., 2015).

우리나라의 경우 연평균 강우량은 세계 평균에 비하여 약 1.6배 많지만, 강우사상의 대부분이 여름철인 6월~9월에 집중되며, 특히 강우의 변동성이 매우 큰 특성으로 인하여 효율적인 수자원관리에 어려움이 있다. 우기기간에 집중호우 및 호우를 동반한 태풍사상으로 인하여 한반도 전역에 많은 강우가 유발되지만 저류시설의 부족 및 대부분의 하천에서는 바다로 유하되는 특성으로 수자원에 상당한 손실을 끼치고 있다. 이러한 수자원 손실을 최소화하고자 대표적인 수공구조물인 댐을 활용한 수자원 확보가 이루어지고 있다. 댐은 하천의 흐름을 차단하거나 흐름의 방향을 바꾸고 늦추는 등의 역할을 하도록 강을 가로질러 세워지는 구조물을 말하는 것으로 우리나라는 용수공급, 발전, 수질조절, 홍수방어 및 가뭄해갈 등과 같은 다양한 목적으로 댐을 활용하고 있다.

효율적인 수자원관리와 유역의 이수 및 치수능력 증대 목적으로 이용되는 대표적 수공구조물인 댐은 가용 수자원을 최적 배분하기 위해 수자원시스템 분석과정에서 가장 공학적으로 검토되어야 하는 요소이다. 이를 위해서는 장기간의 수문자료가 필수적으로 요구되지만 현재 우리나라 댐 중에서 소양강댐(1974년~)을 제외하고는 40년 이상의 수문자료를 보유하고 있는 댐은 존재하지 않는다. 이러한 한계점을 보완하고자 추계학적 모의발생기법을 통하여 해당 유역에 대해서 단기간의 수문계열을 추계학적으로 모형화하고 장기간의 시계열 모의발생을 통하여 이용할 수 있다. 추계학적 모의발생기법은 관측 수문자료의 통계학적 특성을 유지하면서 보다 장기간의 수문자료를 효과적으로 발생시킬 수 있는 확률 통계학적 기법이다.

강우량자료의 경우 관측기간이 유출량 자료에 비하여 장기간의 자료취득이 용이하여 강우모의를 통한 모의강우량을 강우-유출 모형의 입력 자료로 사용하여 장기간의 유출량을 산정하는 연구는 신뢰성 있는 수문분석을 위해서 국내외적으로 다수의 연구가 수행되었다. Kim et al. (2005)은 지구온난화에 따른 잠재적인 효과를 확인하고자 1차 Markov Chain 모형을 통한 강우모의결과를 대청댐 유역에 적용하여 미래 대청댐 유역의 유출량 감소를 예측한 바 있다. Kwon et al. (2011)은 기후변화에 따른 설계홍수량의 변화를 평가하기 위하여 다변량(multivariate) 강우량 상세화 기법의 결과를 장기 강우-유출 모형의 입력 자료로 사용하여 소양강댐 유역의 적용한 결과 설계빈도가 점차 고빈도로 설정될수록 설계홍수량이 기존 계획에 비하여 증가하는 것으로 전망하였다. Kwon et al. (2013)은 불연속 Kernel Pareto 분포를 이용한 다지점 강우모

의기법과 Bayesian HEC-1 모형을 연계하여 대청댐 유역의 불확실성을 고려한 홍수빈도 곡선을 개발하여 적합성을 평가한 결과 불확실성 범위 내에서 기존 방법론들과 유사한 거동을 나타내는 것을 확인하였다.

기존 연구에서는 단순 Markov Chain 모형과 장기유출모형을 연계한 해석이 수행된 바가 있다. 단순 Markov Chain 모형의 경우 평균적인 강우 특성을 재현하는데 우수한 성능을 발휘하지만 극치자료의 생성에서 한계점을 보이고 있다. 이는 유입량 자료를 생성하는 강우 시나리오로서 가뭄 및 홍수와 같은 극치현상을 재현하는데 동일한 문제점을 발생시킬 가능성이 크다. 이와 더불어 기존 연구에서는 강우-유출모형의 불확실성을 체계적으로 고려하지 못하는 단점이 존재한다. 이러한 점에서 본 연구에서는 극치현상 재현에 유리한 강우모의기법과 더불어 강우-유출 모형의 불확실성을 정량적으로 고려할 수 있는 방안을 수립하는데 목적이 있으며, 궁극적으로 강우자료의 표본오차 및 강우-유출과정의 불확실성을 감안한 일단위 유출시나리오를 생성하기 위한 연구를 수행하였다.

본 논문의 구성은 다음과 같다. 2장에서는 강우모의기법과 연계한 댐 유입량 시나리오 생성을 위하여 불연속 Kernel-Pareto 분포형 기반 Markov Chain 모형 기반 강우모의기법의 이론적 배경을 언급하고 SAC-SMA 장기 강우-유출 모형에 대하여 기술하였다. 3장에서는 두 가지 모형을 연계하여 신뢰성 있는 장기간의 수문자료를 확보하고 있는 K-water 다목적 댐인 소양강댐과 대청댐을 대상으로 적용한 연구결과를 통계적 및 시간적 방법을 통하여 제시하였으며 마지막으로 연구결론을 4장에 수록하였다.

2. 연구방법

2.1 불연속 Kernel-Pareto 분포를 이용한 일강수량 모의기법

자료의 발생시간을 고려한 시변성(time-variation)이 있는 관측자료를 순차적으로 나열한 시계열 자료를 근거로 통계학적 특성을 이용하여 추계학적 모형에 적용하여 자료계열을 확장하는 기법들이 여러 수문학 연구에서 개발되었다. 본 연구에서 기존의 2-State Markov Chain 모형 기반의 일강수량 모의기법으로 Markov Chain 모형을 기반으로 한 다수의 연구(Kim et al., 2014; So et al., 2015)에서 제시한 방법론을 요약하여 정리하였다.

Markov Chain 모형의 기본 이론은 특정한 계열을 예측하고자 하는 경우, 해당 계열의 과거특성을 통해 미래를 예측하

는 것으로 Markov Chain 모형의 기본개념은 추계학적으로 시계열자료의 미래 상태($t+1$)를 도출하는 과정에서 과거 ($t-1$)의 상태와 현재(t)의 상태의 상관관계를 사용하는 이론이다. 특정한 두 상황의 바로 이전의 상태가 바로 이후의 상황에만 영향을 주고 그 이전과 이후의 모든 상황에는 전혀 영향을 주지 않는 것으로 정의된다. Markov Chain 모형의 기본개념은 Eq. (1)과 같은 조건부 확률(conditional probability)로 나타낼 수 있다.

$$p(R_{t+1}|R_t, R_{t-1}, \dots, R_1) \quad (1)$$

여기서 R_t 은 현재의 강우상태를 나타내며 R_{t-1}, R_t, \dots, R_1 은 과거의 강우상태를 나타낸다.

일반적으로 강우시계열은 변동성이 크고 수문학적인 지속성이 크지 않은 시계열자료로서 미래강우상태 R_{t+1} 은 강우기간(wet spell)에 작은 계열 상관성이 존재할 수 있으나 일반적으로 3일 이상의 과거 강우 상태와는 독립(independence)이라고 가정할 수 있다(Thompson, 1984). 일강우량 모의를 수행함에 있어 기존 Gamma 분포를 활용하는 경우 일강우량의 통계적 특성을 효과적으로 재현하는 것으로 알려져 있다. 하지만 Gamma 분포의 경우 평균을 중심으로 일강우량을 모의하게 되므로 분포형 양쪽 끝단의 속성인 극소 및 극대강우량 정보를 효과적으로 재현하지 못하는 문제점이 있다. 극치강우량이 빈번히 포함되는 경우에 극치강우량을 효과적으로 재현하지 못할 가능성이 있어 모의강우량을 과소추정할 수 있다. 이에 본 연구에서는 내삽(interpolation)에 우수한 재현능력을 가지는 핵밀도함수(kernel density function) 방법과 극치값(extreme value) 재현이 유리한 Generalized Pareto Distribution (GPD)을 동시에 반영할 수 있는 불연속 Kernel-Pareto 분포형(Piecewise Kernel Pareto Distribution, PKPD)을 일강우량 모의기법에 적용하였다.

2.1.1 핵밀도 함수

비매개변수적 확률밀도 추정기법인 핵밀도 함수는 특정한 분포의 가정없이 관측 자료로부터 확률분포형 유도가 가능한 장점이 있다. Rosenblatt (1956)는 모든 자료가 발생된 각각의 위치에 정규분포와 같은 독립된 핵함수를 부여하는 핵밀도함수 추정법을 개발하였으며, 이는 모든 실수 x 에 대하여 Eq. (2)와 같이 정의된다.

$$f(x) = \frac{1}{nh} \int \frac{1}{d_{j,k}} K\left(\frac{x - X_j}{hd_{j,k}}\right) dx \quad (2)$$

여기서 X_1, X_2, \dots, X_n 은 독립적으로 동일하게 분포된 관측강우량이며 $hd_{j,k}$ 는 양의 변동 광역폭(bandwidth)이다. Eq. (2)와 같은 핵밀도함수는 추정하고자 하는 위치 x 의 거리에 관계없이 동일한 가중치를 적용함으로써 추정된 밀도의 모양이 매끄럽지 못하며 불연속적인 특징이 있다. 따라서 본 연구에서는 이러한 문제를 해결하기 위하여 평활화 핵밀도함수(smooth kernel density function)를 사용하였다.

본 연구에서는 Rule of Thumb를 이용하여 광역폭을 추정하였다. 일반적으로 핵밀도함수의 광역폭 추정시 Rule of Thumb를 활용하는 경우 최적의 광역폭을 추정이 가능하다고 알려져 있다(Hall et al., 1991; Ruppert et al., 1995).

$$\hat{h} = 2.78 \hat{\sigma} n^{-1/5} \quad (3)$$

여기서 n 은 자료의 개수를 나타내며 $\hat{\sigma}$ 는 자료의 표준편차를 의미한다. 일반적으로 핵함수는 Eq. (4)와 같은 3가지 가정을 만족하여야 한다.

$$\text{Mean} : \int K(t) dt = 1 \quad (4a)$$

$$\text{Variance} : \int tK(t) dx = 1 \quad (4b)$$

$$\text{Area} : \int t^2 K(t) dt = \text{constant} \quad (4c)$$

여기서 t 는 $(x - X_j)/h$ 이고, 일반적으로 $t=0$ 에서 최댓값을 가지며 연속적이고 대칭적인 방정식의 형태를 가진다.

핵함수의 종류는 Rectangular, Gaussian, Epanechnikov, Rajagopalan, Cauchy 등이 존재한다(Table 1). 각 핵함수는 사용목적에 따라 적용성을 검토할 필요가 있으며, 본 연구에서는 Gaussian 핵함수를 이용하여 연구를 진행하였다.

Table 1. Types of kernel function

Function Name	Kernel function
Rectangular	$\frac{1}{2}$ for $ t < 1, 0$ if otherwise
Gaussian	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$
Epanechnikov	$\frac{3}{4}\left(1 - \frac{1}{5}t^2\right)/\sqrt{5}$ for $ t < \sqrt{5}$
Rajagopalan	$\frac{3h}{1-4h^2}(1-t^2)$ for $ t \leq 1$
Cauchy	$\frac{1}{\pi(1+t^2)}$

$$K(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{(h \sqrt{(2\pi)})^d} e^{\left(-\frac{1}{2} \left(\frac{x-x_i}{h}\right)^2\right)} \quad (5)$$

2.1.2 Generalized Pareto Distribution

Pareto 분포는 Wakeby 분포의 변형된 형태로 Generalized Pareto Distribution (GPD)의 누적분포함수 $F(x)$ 와 확률밀도함수 $f(x)$ 는 Eq. (6)과 같이 정의된다(Hosking and Wallis, 1987).

$$F(x) = 1 - \left[1 - \frac{k}{\alpha} (x - \epsilon)\right]^{1/k} \quad (6a)$$

$$f(x) = \frac{1}{\alpha} \left[1 - \frac{k}{\alpha} (x - \epsilon)\right]^{1/k-1} \quad (6b)$$

여기서 k 는 형상 매개변수(shape parameter), α 는 규모 매개변수(scale parameter) 및 ϵ 은 절점기준(threshold)이다. $k \leq 0$ 일 조건에서 x 는 $0 \leq x < \infty$ 이고 $k > 0$ 일 조건에는 $0 \leq x \leq \alpha/k$ 이다. 누가분포함수, 매개변수 및 절점기준을 고려하여 극치강우량은 다음 Eq. (7)의 분위값 함수를 통하여 추정된다.

$$x = \epsilon + \frac{\alpha}{k} [1 - (1 - F)^k] \quad (7)$$

본 연구에서는 핵밀도함수와 GPD의 누가확률밀도함수를 동시에 반영할 수 있는 분포형을 활용하여 일강우량 모의를 수행하였다. 즉, 비초과확률이 0.95 이상에서는 극치강우량 재현에 유리한 GPD를 활용하고, 그 외의 구간에서는 핵밀도함수를 활용하게 된다. 이러한 확률분포형의 형태를 불연속 확률밀도함수(piecewise probability density function)라고 하며, 최근에 혼합분포(mixture distribution)와 함께 복합적인 성분을 동시에 모의하는데 있어 수문통계학 분야에서 많이 활용되고 있다.

본 연구에서 사용된 불연속 확률밀도함수의 개념은 Fig. 1과 같다. 불연속 확률밀도함수와 Gamma분포를 적용하여 추정된 일강우량 모의결과를 경험적 누가확률분포와 함께 보여주고 있다. 불연속 확률밀도함수의 경우 0.95 이상의 극대강우량에 대해서는 GPD가 적용되었고 0부터 0.95까지는 핵밀도함수가 적용되었다. 이는 경험적 누가확률분포와 비교결과 불연속 확률밀도함수 누가확률분포의 경우 분포형의 양 끝단에서 경험적인 누가확률밀도함수와 일치되는 모습을 보

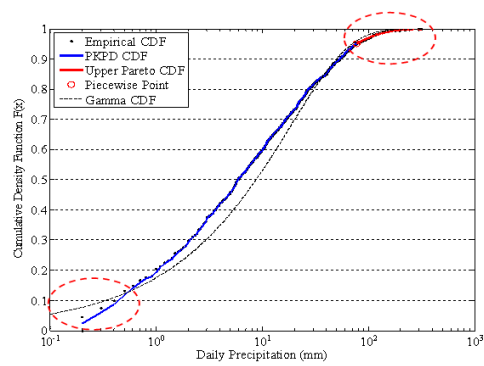


Fig. 1. An example showing the estimated two distinct tails of the piecewise semi parametric CDF for daily rainfall

여주고 있다. 또한, 불연속 확률밀도함수는 비초과확률 0.95 이상에서 원자료가 갖는 극치특성을 매우 잘 묘사하고 있다. 이처럼 분포형에서 Tail의 거동은 극치강우량을 재현하는데 매우 민감한 요소로서 Gamma 분포형은 극치강우량을 모의하는데 무리가 있음을 확인할 수 있다. 이와 더불어 0.05 이하의 누가확률부분에서도 불연속 확률밀도함수는 경험적인 누적분포함수와 유사한 특성을 보이고 있다.

2.2 SAC-SMA 장기유출 모형

Anderson (1973)은 미국 국립기상청(National Weather Service, NWS) 프로젝트를 통하여 강우, 적설과 용설, 토양 함수상태, 하도추적 및 매개변수 최적화 등을 수행할 수 있는 모형을 프로그램으로 구축하여 National Weather Service River Forecast System (NWSRFS)을 개발하였다(Larson et al., 2002). NWS-PC 강우-유출 모형은 NWSRFS의 축소모형으로 토양수분(Soil Moisture)을 고려할 수 있는 SAC-SMA 모형과 HEC-1의 추적모형을 조합하여 홍수 추적시에 운동파(kinematic wave) 추적이나 단위도-Muskinggum 방법을 조합하여 적용할 수 있도록 구축되어 있다(Burnash et al., 1973).

본 연구에서 활용한 유출 모형은 Kim et al. (2015)에 소개된 내용을 정리한 것이다. SAC-SMA 모형은 상부 토양층(upper zone tension water)과 하부 토양층(lower zone tension water)으로 구분되어 상부 토양층은 차단저류지로 직접유출(direct runoff)과 중간유출(inter flow) 과정을 모의하고 하부 토양층은 토양수분함량과 지하수 저류를 모의하게 된다. 상부 토양층과 하부 토양층 사이의 유량의 이동은 수분 이동이 자유로운 자유수 형태와 토양입자의 흡습력에 의한 토양입자에 부착된 부착수 형태로 고려된다. 자유수는 침투와 증발산에 의해 탈수되며, 부착수는 증발산에 의해서만 탈수된다. 이와 같이

Table 2. Description of the parameters of SAC-SMA model

Parameter	Description
UZWWM	Maximum capacity of the upper zone tension storage (mm)
UZFWM	Maximum capacity of the upper zone free water storage (mm)
UZWTC	Upper zone tension water contents (mm)
UZFWC	Upper zone free water contents (mm)
LZWWM	Maximum capacity of the lower zone tension storage (mm)
LZFPM	Maximum capacity of the lower zone primary storage (mm)
LZFSM	Maximum capacity of the lower zone free water storage (mm)
LZTWC	Lower zone tension water contents (mm)
LZFSC	Lower zone free supplemental contents (mm)
LZFPC	Lower zone free primary contents (mm)
ADIMP	Additional impervious area
UZK	Upper zone free water lateral depletion rate
LZPK	Lower zone primary free water depletion rate
LZSK	Lower zone supplemental free water depletion rate
ZPERC	Maximum percolation rate
REXP	Exponent of the percolation equation
PCTIM	Impervious fraction of the watershed area
RESERV	Lower zone free water which cannot be transferred to lower zone tension water
PFREE	Fraction of water percolating from upper zone which goes directly to lower zone free water storage (decimal fraction)

상부 토양층과 하부 토양층을 걸쳐 유출되는 총 하도 유입량은 하도추적이 이루어진 후에 지표하 유출(sub-surface flow)과 유역 출구에서 합산된다. Table 2는 SAC-SMA 모형의 매개변수를 나열한 것으로 상부 토양층에 포함되는 매개변수는 UZWWM, UZFWM, UZK, ADIMP, ZPERC, REXP가 있으며, 하부 토양층에 포함되는 매개변수는 LZTWM, LZFPM, LZFSM, LZSK, LZPK, PFREE가 있다.

2.2.1 Bayesian Markov Chain Monte Carlo 매개변수 최적화
수문모형의 매개변수 최적화를 위해서는 목적함수(object function)를 사용하게 된다. 본 연구에서는 평균 제곱근 오차(Root Mean Square Error, RMSE)를 이용하였다. RMSE는 직관적으로 모형의 정확도를 파악하는 것으로 관측 자료와 모의결과의 차이로 모형을 평가하는 척도로 사용되며 RMSE는 모의결과와 동일 차원을 가지는 지표로서 모의결과가 평균적으로 어느 정도의 오차가 포함되어 있는지를 나타내는 지표로 RMSE 값이 0에 가까울수록 모형성능이 우수하다는 것을

나타낸다. 추가적인 통계적 지표로 관측 자료와 모의결과를 상관계수와 모형의 일치계수(index of agreement, IoA)를 산정하였다. 상관계수를 공분산을 표준편차로 나누어 -1에서 1까지의 값을 가진다. 상관계수는 1 또는 -1에 가까울수록 두 확률변수간의 관련성이 크다는 뜻이고, 0에 가까울수록 관련성이 없다는 것이다. 상관계수와 IoA는 관측 자료의 평균을 사용하여 모의 결과와 관측 자료의 일치성을 나타내는 무차원 지표로서 IoA의 최적값은 1이다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - S_i)^2} \quad (8)$$

$$R = \frac{\sum_{n=1}^N (O_n - \bar{O})(S_n - \bar{S})}{\sqrt{(\sum_{n=1}^N (O_n - \bar{O})^2)(\sum_{n=1}^N (S_n - \bar{S})^2)}} \quad (9)$$

$$IoA = 1 - \frac{\sum_{i=1}^n (O_i - S_i)^2}{\sum_{i=1}^n (|O_i - \bar{O}| + |S_i - \bar{S}|)^2} \quad (10)$$

여기서 O 는 관측자료, \bar{O} 와 \bar{S} 는 관측자료와 모의결과의 평균 값을 의미하며 S 는 모의결과를 의미한다.

위와 같은 통계적 지표를 사용하여 Bayesian Markov Chain Monte Carlo(MCMC) 매개변수 최적화를 수행하였다. Bayesian 이론은 조건부확률(conditional probability)과 우도(likelihood)를 이용하여 사후확률(poster distribution)을 추론하는데 유리하다. Bayesian 이론의 기본개념은 Eq. (11)과 같다.

$$p(\theta, y) = p(\theta)p(y|\theta) \quad (11a)$$

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (11b)$$

$$p(\theta|y) \propto l(\theta|y)p(\theta) \quad (11c)$$

여기서 $p(\theta)$ 는 사전분포로서 우도 $p(y|\theta)$ 와의 곱으로 사후분포 $p(\theta, y)$ 를 추론할 수 있다.

Bayesian 이론에 따른 사전분포와 우도의 곱에 비례하여 모형의 매개변수를 추정하는데 있어서 갱신되는 자료를 바탕으로 우도와 사후분포가 연속적으로 생성되어 이를 통해 모형

매개변수의 불확실성을 정량적으로 분석이 가능하다. 즉, Bayesian 이론은 신규 사후분포가 갱신됨에 따라서 신뢰도를 정정하고자 반복적인 추론을 이어가게 된다(Gelman, 2004).

MCMC는 Markov Chain을 기반으로 한 확률분포로부터 가장한 확률분포를 갖는 표본을 추출하는 알고리즘이다. 충분히 많은 모의 횟수를 거치게 되면 표본은 일반적으로 사용자가 가정한 확률분포에 수렴하게 된다. 다시 설명하면, 다변량 확률분포(multivariate probability distribution)가 복잡하여 이를 따르는 서로 독립(independent and identical distributed, iid)인 난수를 얻을 수 없는 경우에 Markov Chain을 통하여 난수를 사용하여 원하는 분포에 수렴하게 된다. 본 연구에서는 앞 절에서 설명한 Bayesian 이론을 MCMC 기법에 연계하여 보다 빠르게 확률분포의 수렴을 유도하고자 한다.

본 연구에서는 MCMC 기법의 대표적인 방법으로 메트로폴리스 해스팅스(Metropolis-Hasting, MH) 방법을 사용하였다. MH 방법은 Metropolis et al. (1953)이 제안하여 Hastings (1970)에 의해서 일반화 되었다. 각 매개변수의 사후분포 추정 및 매개변수 불확실성 정량화를 수행하기 위해서 본 연구에서는 광범위한 범위에서 매개변수를 탐색하는 전역 최적해(global optimum) 기법을 적용하였다. 이를 통해서 매개변수의 수렴정도를 정량적으로 평가하기 위하여 추적곡선(trace plot)의 수렴조건을 만족하는 경우 최적해로 판단하였다. 본 연구에서 사용된 알고리즘을 서술하면 다음과 같다.

- [1] 사전분포로부터 n 개의 표본 q_1, q_2, \dots, q_n 을 무작위로 추출한 후 Eq. (12)를 이용하여 각 점의 사후분포 $p(q^1|y)$, $p(q^2|y), \dots, p(q^n|y)$ 를 추정한다.

$$likelihood(q^{(t)}|y) = \exp\left[-0.5 \sum_{i=1}^N \left(\frac{\epsilon(q^{(t)})_i}{\sigma}\right)^2\right] \quad (12)$$

- [2] n 개의 점들의 사후확률을 내림차순으로 정리하여 D 에 저장한다. 여기서 D 의 첫 열은 가장 큰 사후확률을 가진 점들의 조합이 된다. 병렬적인 연속점 P_1, P_2, \dots, P_q 을 초기화한다.
- [3] D 의 점들은 각 m 개의 자료를 갖는 q 개의 부분 C_1, C_2, \dots, C_q 로 나누어진다.
- [4] 각 병렬적인 연속 점들은 메트로폴리스 해스팅스 기법에 의해서 연속적으로 갱신되어 사후확률을 추정하게 된다.
- [5] q 개의 부분 C_1, C_2, \dots, C_q 다시 D 공간으로 보내져 사후 확률에 따라 다시 내림차순으로 정리된다. 이를 다시 [3]과 같이 재구성하는 과정을 진행하게 된다.

3. 연구결과

3.1 대상유역

우리나라 댐은 용도에 따라 다목적 댐, 홍수조절 댐, 생활·공업용수 댐, 수력발전 댐 및 농업용수 댐으로 구분된다. 본 연구에서는 수계의 대표성을 가지면서 장기간의 관측 자료가 확보된 한강권역 소양강댐과 금강권역 대청댐을 대상유역으로 설정하였다.

소양강댐은 이수 및 치수 측면에서 상당한 기여를 하고 있는 우리나라 최대 규모의 다목적 댐으로, 1974년 관측개시를 하였으며 유역면적은 2,703 km²로서 강원도 인제군 외 4개 시군(춘천, 양구, 고성, 홍천)에 걸쳐있다. 춘천시에서 동북쪽으로 13 km, 북한강 합류지점에서 12 km 떨어진 소양강 계곡에 위치한 높이 123 m, 길이 530 m, 체적약 9,600,000 m³의 사력댐이다. 총 저수용량은 29억 m³이며, 홍수조절용량 5억 m³을 보유, 발전시설용량 10만 kW 2기를 설치하여 운영 중이다. 대청댐은 금강수계의 수자원을 다목적으로 개발하기 위해 설치된 댐으로 1981년 관측개시를 하였으며 유역면적 3,204 km², 유로 총연장 262.7 km이며 금강 중류부에 높이 72 m, 길이 495 m, 체적 1,234,000 m³의 콘크리트 중력식 댐과 석괴식댐의 복합형 댐이다.

3.2 PKPD 강우모의결과

장기 강우-유출 모형의 입력 자료로 활용하기 위한 댐 유역 강우량을 추계학적으로 모의하는데 있어서 핵심적인 사항은 관측강우계열의 연속성 및 통계학적 특성을 효과적으로 반영하는지 여부이다. 본 연구에서는 장기유출 모형의 입력 자료로 사용되는 모의강우계열의 연속성을 평가하기 위하여 연 최대 무강우일 발생사상을 Fig. 2에 도시하였다. Fig. 2를 살펴보면 PKPD를 통하여 모의된 강우계열은 전반적으로 관측강우계열의 연 최대 무강우 패턴을 효과적으로 인지하고 있음을 확인할 수 있다.

추가적으로 PKPD 강우모의 기법을 통한 모의강우의 통계적 특성을 확인하기 위하여 강우계열의 평균적인 특성뿐만 아니라 표준편차, 왜곡도 및 첨예도를 산정하였다.

분석대상 댐의 통계적 특성을 살펴보면 관측치의 통계특성을 매우 효과적으로 재현하는 것을 Fig. 3을 통하여 확인할 수 있다. 이는 모의강우량을 장기 강우-유출 모형의 입력 자료로 사용함에 있어 적합한 결과로 사료된다. 유역의 유출현상은 강우에 가장 민감하게 반응하기 때문에 강우사상의 양적인 특성뿐만 아니라 강우사상의 발생여부는 유출분석을 수행함에 있어 면밀히 검토되어야 하는 부분이다. Table 3은 PKPD

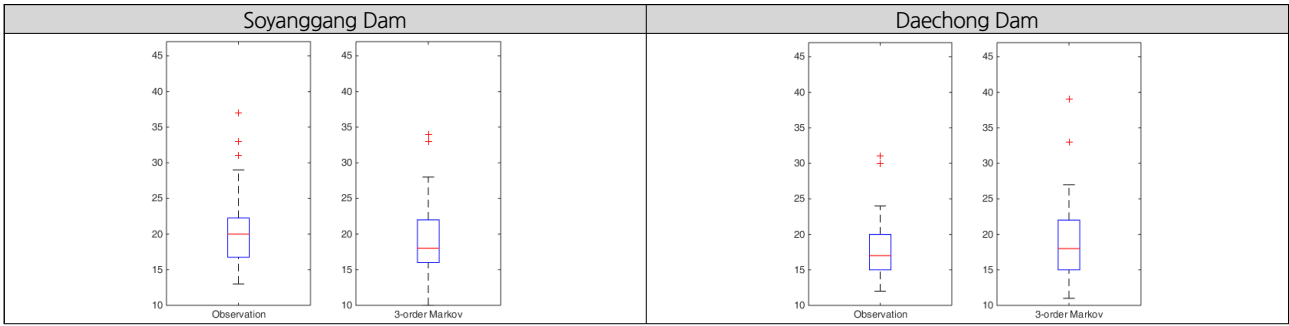


Fig. 2. Maximum dry spell for observation and simulation

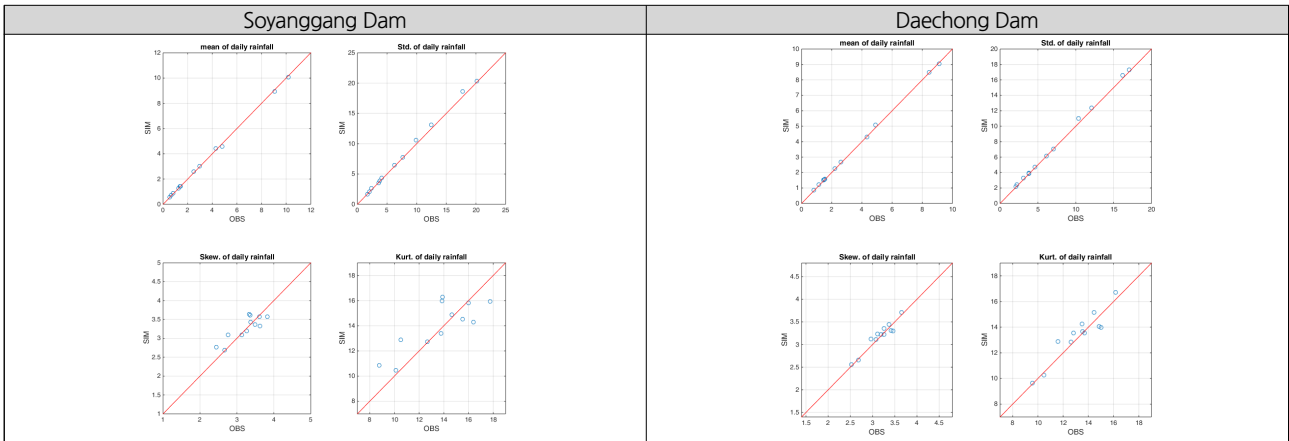


Fig. 3. Scatter plot of static moments for PKPD rainfall simulator

Table 3. Comparison of PKPD rainfall simulation result (occurrence)

		Maximum Rainfall Occurrence	Minimum Rainfall Occurrence	Average Rainfall Occurrence
Soyonggang Dam	Observation	149	95	122
	PKPD	151	93	122
Daechong Dam	Observation	158	100	129
	PKPD	149	106	129

강우모의 결과의 강우발생 특성을 비교한 결과이다. 평균 강우일수의 경우 분석대상 댐의 관측강우발생 특성과 일치하는 결과를 확인하였으며, 최대 강우일수와 최저 강우일수는 관측 강우발생일과 다소 차이는 있지만 전반적으로 관측 강우발생 특성을 효과적으로 재현하고 있음을 확인하였다(So, 2012).

3.3 SAC-SMA 모형 유입량 산정결과

SAC-SMA 모형의 매개변수 최적화를 위한 Bayesian MCMC 모의횟수는 20,000회 수행하였다. 일반적으로 강우-유출 모형의 예측성은 매개변수의 적합성에 의존하여 예측

결과가 목적함수를 만족하는 경우 유효한 매개변수로 인지하게 된다. 본 연구에서는 Bayesian MCMC 모의를 통하여 산정된 매개변수 중에서 안정적으로 수렴이 이루어지고 목적함수를 효과적으로 만족하는 매개변수 사상을 최적 매개변수로 설정하였다. 본 연구에서 적용한 Bayesian MCMC 기법은 매개변수 최적화 및 불확실성 정량화를 위하여 다수의 연구에서 적용되어 우수한 모형성능이 확인되었으며(Kwon et al., 2008; Lima and Lall, 2010; Russo et al., 2015) 이는 불확실성을 정량적으로 고려하기 위한 본 연구의 목적에 부합하는 것으로 사료된다. Fig. 4는 최적매개변수를 대상으로 일단위 수문곡선을 도시한 결과로 파란색 ★는 관측유량을 의미하고 검은색 실선은 검증유량을 의미한다. 분석대상 댐에서 관측된 극치유량을 효과적으로 모의하고 있는 것을 확인할 수 있으며, 이는 SAC-SMA 강우-유출 모형이 극치사상을 효과적으로 인지하고 있으며 Bayesian MCMC로 산정된 매개변수가 적합성을 가지고 있는 것으로 사료된다.

가장 모의성능이 우수한 매개변수를 활용하여 댐 별로 월 주기 수문곡선을 Fig. 5에 도시하였다. ●은 중간값(median)을 의미하며 실선은 평균값(mean)을 의미한다. 연구대상 댐

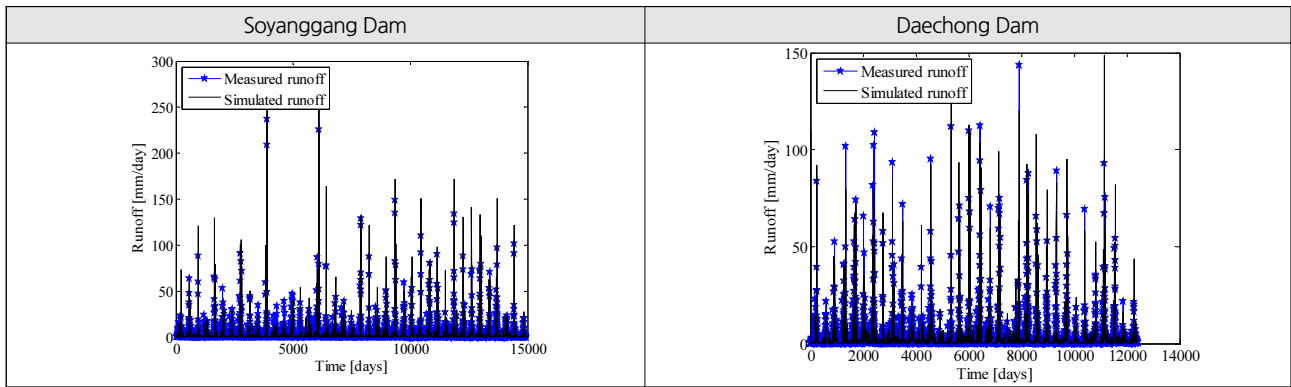


Fig. 4. Hydrographs using best parameter set for daily inflow simulation

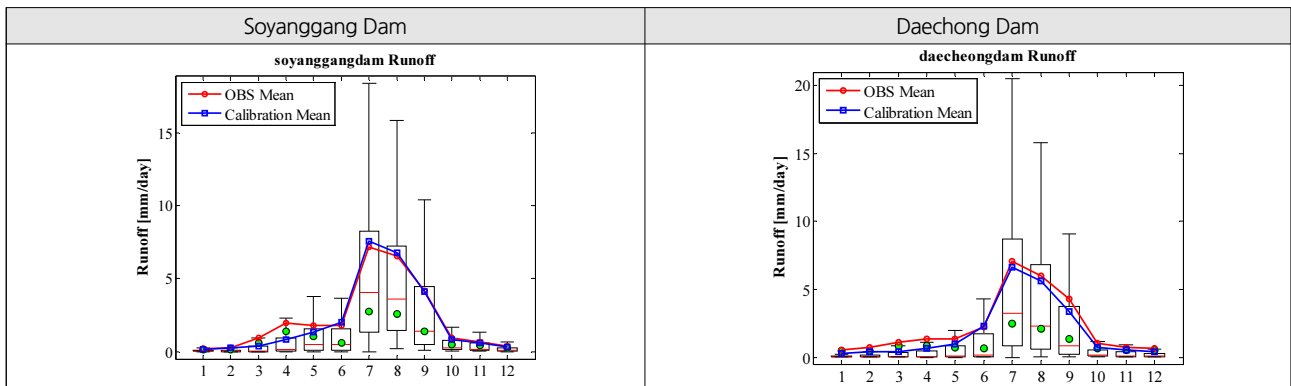


Fig. 5. Hydrographs using best parameter set for monthly inflow simulation

Table 4. Model calibration statistical results based on bayesian MCMC

		Mean (mm/day)	Standard deviation	Skewness	R	IoA	RMSE (mm/day)
Soyganggang Dam	Observation	2.250	7.341	12.632	0.875	0.932	0.153
	SAC-SMA	2.118	7.911	14.744			
Daechong Dam	Observation	2.269	6.539	8.740	0.817	0.890	0.498
	SAC-SMA	2.099	5.460	10.265			

의 경우 3월 및 4월에 일부 왜곡현상이 발생하지만 전반적으로 홍수기 및 비홍수기 유량을 관측유량과 유사하게 모의하고 있는 것을 확인할 수 있다. 연구대상 댐의 경우 관측유량의 월별 유량패턴도 효과적으로 관측유량의 패턴을 나타내고 있는 것을 확인할 수 있다. Table 4는 통계적 모멘트와 통계적 지표를 산정한 결과이다. Bayesian MCMC를 통하여 산정된 매개변수는 상관계수의 경우 0.82~0.88을 나타내고 있으며, IoA의 경우 0.89~0.93으로 높은 적합성을 확보하고 있다. 통계적 모멘트 역시 1차 모멘트에서부터 3차 모멘트까지 관측자료의 통계적 특성을 효과적으로 재현한 것을 확인하였다.

Fig. 6은 Fig. 5와는 다르게 모의된 강우와 Bayesian MCMC 과정으로 추정된 사후분포를 활용하여 유출시나리

오를 도출하였다. 즉, 일단위 100년 기간의 100개의 PKPD 강우모의 시나리오와 SAC-SMA 유출모형의 매개변수 산정결과를 통하여 산정된 최적 매개변수의 사후분포를 입력 자료로 구축하여 다수의 댐의 일유입량 시나리오를 산정하여 월주기 수문곡선을 도출하였다. ●은 중간값을 의미하며 실선은 평균값을 의미한다. 도시결과를 살펴보면 평균값의 경우 관측 자료를 전반적으로 유사하게 모의하는 것을 확인할 수 있으며, 중간값의 경우 Boxplot의 신뢰구간에 효과적으로 위치하고 있는 것을 확인할 수 있다. 소양강댐의 경우 7월의 모의유량은 다소 과소추정되는 경향이 있지만 대청댐의 경우 홍수기와 비홍수기의 모의유량이 관측유량을 매우 유사하게 모의하는 결과를 확인할 수 있다.

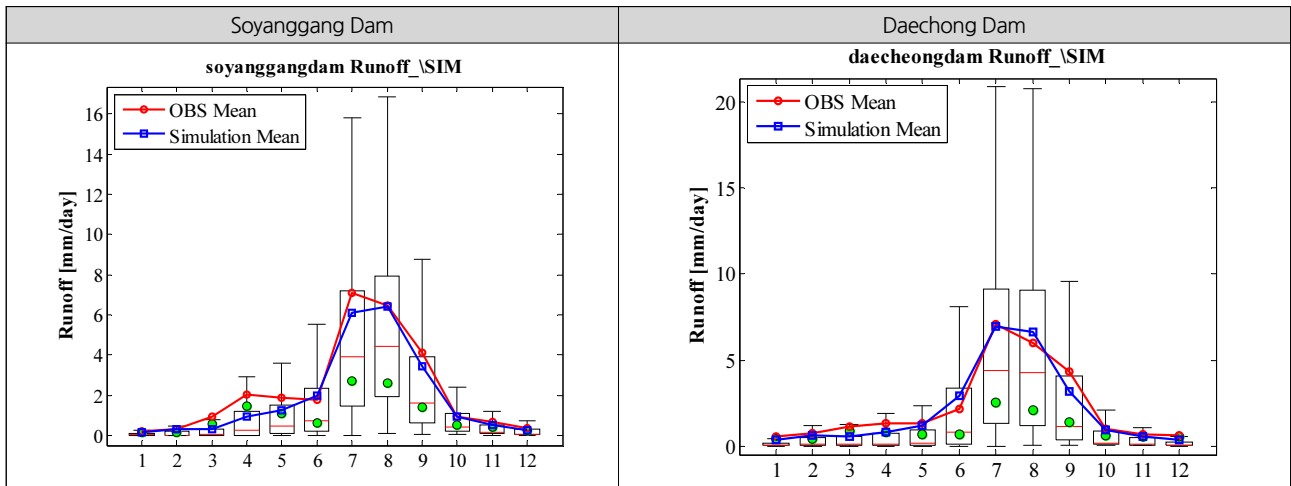


Fig. 6. Monthly hydrographs using posterior parameter distribution model parameters

Table 5. Scenario based reliability analysis for water supply

		Annual			Daily		
		Total	Deficiency	%	Total	Deficiency	%
Soyonggang Dam	Observation	41	0	100	14,965	0	100
	Simulation	100	2	98	36,500	52	99.9
Daechong Dam	Observation	34	6	82.4	12,410	426	96.6
	Simulation	100	10	90.0	36,500	121	99.7

3.4 용수공급능력 평가

용수수요는 그 사용목적에 따라 생활용수, 공업용수, 농업용수 및 하천유지용수 등으로 구분되며, 각 목적별 용수에 대하여 목표연도에 따른 단기수요와 장기수요로 구분된다. 본 연구결과를 보다 실용적으로 사용하기 위해서는 시나리오 기반의 장기 용수공급능력 평가를 수행하였다. 소양강댐의 경우 일별 생활용수 공급량은 $38.1 \times 10^6 \text{ m}^3$ 및 유지용수 $8.10 \times 10^6 \text{ m}^3$ 를 공급하며 월별로 $111.8 \sim 126.4 \times 10^6 \text{ m}^3$ 를 공급하고 있다. 대청댐은 일별 생활용수 공급량은 $41.2 \times 10^6 \text{ m}^3$ 및 6월에 농업용수 최대량 $47.6 \times 10^6 \text{ m}^3$ 를 공급하며 월별로 $99.7 \sim 230.2 \times 10^6 \text{ m}^3$ 를 공급하고 있다. 본 연구에서는 제시된 용수공급량과 도출된 댐별 일유입량 시나리오를 활용하여 현재 용수공급 조건에 따른 용수공급능력을 평가하였다. Table 5는 100개의 유입량 시나리오를 통하여 100개의 용수공급능력 시나리오를 평가한 결과를 나타내며, 소양강댐은 부족연수 2회(신뢰도 98%), 부족일수 52회(신뢰도 99.9%)로 나타났으며 대청댐의 경우 부족연수 10회(신뢰도 90.0%), 부족일수 121회(신뢰도 99.7%)로 나타났다. 대청댐의 경우 관측기간(1981년~2014년)에 비하여 개선된 모의결과를 확인하였다.

4. 결론

현재 세계적으로 기후변화로 인한 이상기후로 인하여 기상관련 재해가 과거에 경험하지 못했던 규모로 빈번하게 일어나고 있다. 한반도의 경우 국토의 대부분이 산악지형(약 70%)인 특성으로 인해 지형학적 특성에 큰 영향을 받는 기상현상의 공간적 편차가 상당히 크다. 이러한 점에서 우리나라는 기후변동성과 지형학적 영향이 맞물려 지역적으로 불균형하게 발생하는 강우사상으로 효율적인 수자원관리가 매우 어려운 실정이다. 기후변화로 인한 미래 수자원 확보 및 수자원 관리를 위해서는 기존에 관측된 수문자료를 활용하여 수문 통계학적 모의발생기법을 통하여 모의수문변량을 생산할 수 있다. 이렇게 생산된 모의수문변량을 향후 미래에 발생 가능한 수문사상으로 가정하여 미래 수자원계획 및 의사결정을 위한 핵심적인 정보로 제공될 수 있다.

수자원관리에 있어 핵심적인 고려사항인 댐은 국내의 경우 동아시아 주변국인 중국과 일본에 비하여 숫자가 적어 소수의 댐으로 우기기간에 발생하는 강우를 저류하여 이듬해 우기기간까지 농업용수, 생활용수, 공업용수 및 발전용수 등 다양한 목적으로 활용하고 있다. 하지만 최근 가뭄현상과 연

계하여 수자원 부족에 대한 심각성이 부각되고 있지만 환경문제 및 경제적 문제로 인하여 기존 댐을 효율적으로 활용하는 대안에 무게가 실리고 있다. 따라서 본 연구에서는 기존 댐을 활용한 수자원관리를 위해서 강우모의기법과 강우-유출 모형을 연계한 댐 일유입량 산정연구를 수행하였다. 이를 위해서 기존 Markov Chain 모형에서 일강우량의 평균적인 특성과 극치특성을 동시에 고려할 수 있는 PKPD 강우모의기법과 세계적으로 널리 사용되고 있는 SAC-SMA 강우-유출 모형에 Bayesian Markov Chain Monte Carlo 기법을 연계하여 19개 매개변수를 최적화하고 각 매개변수의 사후분포를 도출하여 소양강댐과 대청댐의 일유입량을 생성하였다. 본 연구를 수행하여 얻은 결론은 다음과 같다.

- 1) 강우-유출 모형의 입력 자료로 활용하기 위한 PKPD 강우모의기법은 극치강우량을 일반강우량으로부터 분리하여 모의함으로써 기존 강우모의기법에 내재된 문제점을 해결하였으며, 낮은 차수의 통계적 모멘트를 정확하게 재현할 수 있는 부가적인 장점을 확인하였다. 이는 여름철 극치강우량을 재현하는데 유리한 기법으로 사료되어 유출모의시 극치유입량을 효과적으로 재현할 가능성을 확보한 것으로 판단된다.
- 2) 강우-유출 매개변수 최적화를 위하여 사용한 Bayesian MCMC 기법은 안정된 매개변수 산정을 가능케 하였으며, 매개변수의 사후분포를 이용한 모의유출량은 관측유량의 통계적 특성을 효과적으로 모의하는 것을 확인하였다. 이는 향후 해당유역의 수자원계획을 수립함에 있어 모의유량 시나리오를 활용한 불확실성 분석이 가능할 것으로 판단된다.
- 3) PKPD 모형과 SAC-SMA 강우-유출 모형은 집중호우와 극한홍수와 같은 고빈도 강우량과 홍수량을 동시에 재현할 수 있는 모형성능을 확인하였으며, 이와 같은 결과는 기후변화에 따른 극한 수문사상을 효과적으로 인지할 수 있는 수문모형으로 판단된다.
- 4) 본 연구를 통하여 산정된 모의유량은 통계적으로 관측 자료를 효과적으로 재현하고 있기 때문에 모의유량을 활용하여 수자원계획 뿐만 아니라 유출량을 입력 자료로 활용할 수 있는 발전용량 산정, 용수공급 시나리오 및 수질 예측 모니터링 등 다양한 분야에서 활용이 가능할 것으로 판단된다.

향후 연구로는 본 연구에서 구축된 연계모형에 기후변화 시나리오를 추가적으로 반영한다면 각 댐별 일유입량을 정량

적으로 산정하여 권역별·수계별 수자원관리를 위한 기초자료로 활용성이 클 것으로 사료된다.

감사의 글

본 연구는 국토교통부 물관리연구사업의 연구비지원(15AWMP-B079625-02)에 의해 수행되었습니다. 저자들은 전북대학교 방재연구센터에 소속되어 연구를 수행하였습니다.

References

- Anderson, E.A. (1973). "National Weather Service river forecast system : Snow accumulation and ablation model." US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service.
- Burnash, R.J.C., Ferral, R.L., and McGuire, R.A. (1973). "A generalized streamflow simulation system, conceptual modeling for digital computers." Joint Federal, State River Forecast Center, Sacramento, CA.
- Diffenbaugh, N.S., Swain, D.L., and Touma, D. (2015). "Anthropogenic warming has increased drought risk in California." *Proceedings of the National Academy of Sciences*, Vol. 112, No. 13, pp. 3931-3936.
- Fallah, B., and Cubasch, U. (2015). "A comparison of model simulations of Asian mega-droughts during the past millenium with proxy reconstructions." *Climate of the Past*, Vol. 11, No. 2, pp. 253-263.
- Gelman, R. (2004). "Cognitive development. Stevens." *handbook of experimental psychology*.
- Hall, P., Sheather, S.J., Jones, M.C., and Marron, J.S. (1991). "On optimal data-based bandwidth selection in kernel density estimation." *Biometrika*, Vol. 78, No. 2, pp. 263-269.
- Hastings, W.K. (1970). "Monte Carlo sampling methods using Markov chains and their applications." *Biometrika*, Vol. 57, No. 1, pp. 97-109.
- Hosking, J.R., and Wallis, J.R. (1987). "Parameter and quantile estimation for the generalized Pareto distribution." *Technometrics*, Vol. 29, No. 3, pp. 339-349.
- Kim, Y.O., Seo, Y.W., Lee, D.R., and Yoo, C.S. (2005). "Potential Effects of global warming on a water resources system in Korea." *Water International*, Vol. 30, No. 3, pp. 400-405.
- Kim, T.J., Kwon, H.H., Lee, D.Y., and Yoon, S.K. (2014). "Development of Stochastic Downscaling Method for Rainfall Data Using GCM" *Journal of Korea Water Resources Association*, Vol. 47, No. 9, pp. 825-838.
- Kim, T.J., Jeong, G.I., Kim, K.Y., and Kwon, H.H. (2015). "A

- Study on Regionalization of Parameters for Sacramento Continuous Rainfall-Runoff Model Using Watershed Characteristics." *Journal of Korea Water Resources Association*, Vol. 48, No. 10, pp. 793-806.
- Kwon, H.H., Kim, J.G., and Park, S.H. (2013). "Derivation of Flood Frequency Curve with Uncertainty of Rainfall and Rainfall-Runoff Model." *Journal of Korea Water Resources Association*, Vol. 46, No. 3, pp. 59-71.
- Kwon, H.H., Sivakumar, B., Moon, Y.I., and Kim, B.S. (2011). "Assessment of change in design flood frequency under climate change using a multivariate downscaling model and a precipitation-runoff model." *Stochastic Environmental Research and Risk Assessment*, Vol. 25, No. 4, pp. 567-581.
- Kwon, H.H., Brown, C., and Lall, U. (2008). "Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling." *Geophysical Research Letters*, Vol. 35, No. 5. DOI: 10.1029/2007GL032220.
- Larson, L., Singh, V.P., and Frevert, D. (2002). "National Weather Service River Forecast System (NWSRFS)." *Mathematical models of small watershed hydrology and applications* pp. 657-703.
- Lima, C.H., and Lall, U. (2010). "Spatial scaling in a changing climate: A hierarchical bayesian model for non-stationary multi-site annual maximum and monthly streamflow." *Journal of Hydrology*, Vol. 383, No. 3, pp. 307-318.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. (1953). "Equation of state calculations by fast computing machines." *The journal of chemical physics*, Vol. 21, No. 6, pp. 1087-1092.
- Rosenblatt, M. (1956). "Remark on some nonparametric estimates of a density function." *The Annals of Mathematical Statistics*. Vol. 27, No. 3, pp. 832-837.
- Ruppert, D., Sheather, S.J., and Wand, M.P. (1995). "An effective bandwidth selector for local least squares regression." *Journal of the American Statistical Association*, Vol. 90, No. 432, pp. 1257-1270.
- So, B.J., Kwon, H.H., Kim, D.K., and Lee, S.O. (2015). "Modeling of daily rainfall sequence and extremes based on a semiparametric Pareto tail approach at multiple locations." *Journal of Hydrology*, Vol. 529, pp. 1442-1540.
- So, B.J. Development of Multisite Daily Rainfall Simulation Model Using Piecewise Kernel-Pareto Continuous Distribution. Master's Thesis, Chonbuk National University, Jeonju, Jeollabuk, Republic of Korea.
- Russo, T.A., Devineni, N., and Lall, U. (2015). "Assessment of Agricultural Water Management in Punjab, India, Using Bayesian Methods." In *Sustainability of Integrated Water Resources Management*, Springer International Publishing, pp. 147-162.
- Thompson, C.S. (1984). "Homogeneity analysis of rainfall series: an application of the use of a realistic rainfall model." *Journal of Climatology*, Vol. 4, No. 6, pp. 609-619.
- Wang, S.Y., Hipps, L., Gillies, R.R., and Yoon, J.H. (2014). "Probable causes of the abnormal ridge accompanying the 2013–2014 California drought: ENSO precursor and anthropogenic warming footprint." *Geophysical Research Letters*, Vol. 41, No. 9, pp. 3220-3226.